Spatio-temporal organisation of replication

Part I : Evidence of large scale gradients in the orientation of fork progression

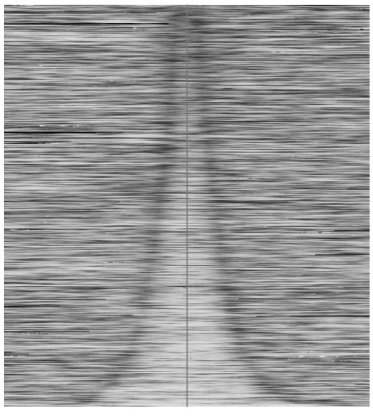
Chun-Long Chen, Yves d'Aubenton-Carafa & Claude Thermes
Centre de Génétique Moléculaire (CNRS), Allée de la Terrasse, Gif-sur-Yvette, France

Guillaume Guilbaud, Aurélien Rappailles & Olivier Hyrien
Eukaryotic chromosome replication group, Ecole Normale Supérieure, CNRS, Paris, France

Arach Goldar
Commissariat à l'Energie Atomique (CEA), iBiTec-S, Gif-sur-Yvette

Benjamin Audit, Antoine Baker, Antoine Leleu, Lamia Zaghloul & Alain Arneodo
Laboratoire Joliot-Curie et Laboratoire de Physique, Ecole Normale Supérieure de Lyon, CNRS, Lyon, France

Human DNA replication initiates from a large number of origins according to a spatio-temporal program depending of the tissue and/or the development stage. Understanding this program is a major issue but few origins have still been characterized. Origin specification depends on still poorly known chromatin features rather than on recognition of simple DNA sequence motifs. We have shown by bioinformatic analyses that a third of the human genome is organized in "N-domains", i.e. ~1Mb segments whose borders contain potential replication origins and which show a characteristic N-shaped profile of the nucleotide compositional skew[1]. We have determined the replication timing profile of the entire human genome by massive sequencing of replication intermediates[2] and analysis of this profile allows us to establish that borders of the N-domains correspond to efficient replication origins. We also demonstrate that the N-like pattern of the skew profile arises from an asymmetry of the mutation/repair processes between the leading and lagging replicating strands and that it reflects the mean polarity of the replicating forks. Detailed analysis of the data suggests a gradient of replication fork orientation that spreads from the borders to the center of N-domains: replication first initiates at master origins located at N-domain borders followed by a cascade of less efficient, secondary initiations within N-domains. The results will be discussed in the evolutionary context of genome organization.



Each line represents the replication timing profile centered at the middle of each N-domain; the N-domains are ranked by increasing size from top to bottom; dark (clear) grey corresponds to early (late) replication timing values.

1. Brodie of Brodie EB, *et al.* (2005) From DNA sequence analysis to modeling replication in the human genome. *Phys. Rev. Lett.* 94:248103. Touchon M, *et al.* (2005) Replication-associated strand asymmetries in mammalian genomes: Toward detection of replication origins. *Proc. Natl. Acad. Sci. USA* 102:9836-9841. Huvet M, *et al.* (2007) Human gene organization driven by the coordination of replication and transcription. *Genome Research* 17:1278-1285.
2. Chen CL, *et al.* (2010) Impact of replication timing on non-CpG and CpG substitution rates in mammalian genomes. *Genome research* 20:447-457.